geographical analysis

Geographical Analysis (2014) **46**, 37–52

Modeling Net Ecosystem Carbon Dioxide Exchange Using Temporal Neural Networks after Wavelet Denoising

Fatih Evrendilek

Department of Environmental Engineering, Abant Izzet Baysal University, Bolu, Turkey

Eddy covariance (EC) time-series data obtained from flux towers are noisy due to both stochastic atmospheric turbulence and deterministic processes, and no standard data-denoising protocols exist for them. The potential of six temporal artificial neural networks (ANNs) augmented with and without three orthogonal wavelet functions was tested for predicting net ecosystem exchange of carbon dioxide (CO_2) based on a long-term EC data set for a temperate peatland. Multiple comparisons were made of (1) temporal ANNs with and without discrete wavelet transform (DWT) denoising; (2) denoising with the orthogonal wavelet families of Daubechies, Coiflet, and Symmlet; (3) different decomposition levels; (4) time-delay neural network, time-lag recurrent network, and recurrent neural network; (5) online learning versus batch learning algorithms; and (6) diel, diurnal, and nocturnal periods. The coefficient of determination, root mean square error, and mean absolute error performance metrics were used for multiple comparisons based on training, cross-validation, and independent validation of the temporal ANNs as a function of 24 explanatory variables contained in an EC data set. Integration of the temporal ANNs and DWT denoising provided more accurate and precise estimates of net ecosystem CO_2 exchange.

Introduction

Exchanges and budgets of carbon dioxide (CO_2), water (H_2O), and energy at ecosystem scales are commonly quantified using eddy covariance (EC) systems to better understand local and global spatiotemporal changes in the climate system. A pressing need to mitigate and monitor effects of global climate change led to the establishment of a global network of over 500 EC flux towers (FLUXNET) that continuously directly measure at a sampling rate of 5–50 Hz across different ecosystem types (Baldocchi 2008). FLUXNET data enable us to predict spatiotemporal dynamics of net CO_2, H_2O, and energy exchanges among the pedosphere, hydrosphere, biosphere, and atmosphere, and to upscale flux patterns from an ecosystem scale to a global scale in a changing

Correspondence: Fatih Evrendilek, Department of Environmental Engineering, Abant Izzet Baysal University, Golkoy Campus, 14280 Bolu, Turkey
e-mail: fevrendilek@ibu.edu.tr

Submitted: March 2, 2012. Revised version accepted: March 25, 2013.

doi: 10.1111/gean.12025

global environment. Limitations of in situ EC sensors, the stochastic nature of atmospheric turbulent eddies, nonlinear interactions of complex environmental processes, and the lack of standard EC data-processing protocols contribute to the creation of stochastic and deterministic noises contained in EC data.

The literature reports developments and implementations of artificial neural networks (ANNs) to EC data (e.g., van Wijk and Bouten 1999; van Wijk, Bouten, and Verstraten 2002; Papale and Valentini 2003; Melesse and Hanley 2005; Moffat et al. 2007; Jung et al. 2011). Van Wijk and Bouten (1999) apply ANNs to select the best predictors to model the H_2O and CO_2 exchange of coniferous forest ecosystems. Van Wijk, Bouten, and Verstraten (2002) compare the efficacy of ANNs, fuzzy logic, an index model, and a process-based model to obtain insights into climate change effects, parameterization, and upscaling of forest responses to a regional scale. Based on EC measurements of CO_2 fluxes at 16 European sites, Papale and Valentini (2003) show that ANNs are useful for gap-filling procedures, the detection of outliers, and the improvement of parameterization in atmospheric and biospheric models. In comparison with other gap-filling strategies, ANNs appear to replicate underlying patterns in the EC data with the highest R^2 and the lowest root mean square error (RMSE), particularly, during daytime, while being more sensitive to small changes in meteorological inputs (Melesse and Hanley 2005; Moffat et al. 2007). Jung et al. (2011) upscale the local-scale FLUXNET observations of CO_2, H_2O, and energy fluxes to the global scale using a machine learning technique (i.e., model tree ensembles).

In addition, numerous statistical techniques applied to EC data include time-series analysis, regression analysis, look-up tables, and stochastic gap-filling procedures (Falge et al. 2001; Hollinger et al. 2004; Hui et al. 2004; Knorr and Kattge 2005; Reichstein et al. 2005; Gove and Hollinger 2006; Ooba et al. 2006; Richardson et al. 2007; Stauch, Jarvis, and Schulz 2008). To date, a few wavelet transform algorithms have been used to analyze EC data (Braswell et al. 2005; Kaheil et al. 2008; Saito and Asanuma 2008). The Haar (1910), Daubechies (1988), Coiflet (Daubechies 1992), and Symmlet (Daubechies 1992) orthogonal wavelet families have been adopted mostly for high-frequency data because they enable any signal to be decomposed and the decomposed signal to be reconstructed without losing information (Mallat 1999). For example, Braswell et al. (2005) select the orthogonal wavelet transform with the wavelet basis of Coiflet 5 (coif5) for the decomposition of EC time series into multiple time scales, from diurnal to decadal, exploiting its symmetrical form that maintains the characteristic shape of the diurnal and seasonal cycles, and assess model performance for the multiple time scales. Kaheil et al. (2008) use the orthogonal wavelet transform with the Haar function to downscale concurrently coarse-resolution satellite data of evapotranpiration and to forecast short-term evapotranpiration at a finer spatial resolution. Saito and Asanuma (2008) demonstrate the capability of the orthogonal wavelet transform with the Haar basis to detect the spectral gaps that separate turbulent and mesoscale motions of sensible heat and water vapor fluxes measured over a rice field and a larch forest. However, these preceding studies focus either on gap-filling techniques or on reduction of uncertainties, without the integration of temporal ANNs and wavelet denoising.

The temporal ANNs are better suited to modeling nonlinear dynamic patterns embedded in time-series EC data and include time-delay neural networks (TDNN), time-lag recurrent networks (TLRN), and recurrent neural networks (RNN). The temporal ANNs involve short- or long-term memory structures that provide a limited temporal history of events or delayed feedback (recurrent) loops, respectively. The TDNN model has high memory resolution (how

accurately information about the number of input variables per a given time unit is maintained) but low memory depth (how far into the past the memory extends for information in relation to the memory size), whereas the TLRN and RNN models have low memory resolution but high memory depth (e.g., Fuchs et al. 2009). In contrast with the RNN models, which are very difficult to train due to their rich dynamic behavior and sensitivity to deviations, the TLRN models are characterized by their stability and low sensitivity to noise (e.g., Sattari, Yurekli, and Pal 2012). Batch learning (B) uses multiple passes (known as epochs) repeatedly processing an entire previously used training set plus new examples and considers all the training instances at once (Oza 2005). Online learning (O) (i.e., sample-by-sample learning) uses only one epoch through the entire training set without processing the previously learned data and considers one training instance individually at a time (Oza 2005). Unlike online learning, the performance and applicability of batch learning generally are limited by how much training data a given system memory can handle. However, online learning cannot fully optimize a cost function defined for training examples, whereas batch learning can completely optimize a cost function defined for training examples (Nguyen, Bouzerdoum, and Phung 2008).

A lack of studies exists about the use of discrete wavelet transform (DWT) denoising and temporal partitioning in temporal ANNs to predict EC-based net ecosystem exchange (NEE) of CO_2 for a peatland. Also, no information exists about the comparative performances of the combined uses of temporal ANNs and DWT denoising with the different orthogonal families. The objectives of this article are as follows: (1) to compare performances of temporal ANNs with and without DWT denoising and temporal partitioning (diurnal, nocturnal, and diel), and (2) to identify the best combination of the six temporal ANN algorithms of TDNN-1-O-M and -1-B-M, RNN-1-O-M and -1-B-M, and TLRN-1-O-M and -1-B-M (1 denotes the number of hidden layers, B denotes batch learning, O denotes online learning, and M denotes momentum algorithm) and the three orthogonal wavelet families of Coiflet, Symmlet, and Daubechies with an order of 10 to find the best method to model an approximation of the EC-based data at multiple time scales.

Data collection

The net turbulent flux densities of CO_2 (F_c, mg m^{-2} s^{-1}) above the vegetation canopy of a temperate peatland in Yenicaga (Bolu, Turkey) were measured between day of year (DOY) 203 (7 July 2010 11:00) and DOY 719 (12 December 2011 19:00) using an EC system. The EC system mounted on a 3-m tower used in this study consists of an open-path CO_2/H_2O gas analyzer (LI-7500, LI-COR Biosciences, Lincoln, NE, USA), a three-dimensional (3-D) sonic anemometer/thermometer (CSAT3, Campbell Scientific, Inc., Logan, UT, USA), and a data logger (CR3000, Campbell Scientific, Inc.). CO_2 fluxes and 24 ancillary associated measurements (see Table 1) were recorded at a sampling rate of 10 Hz, block-averaged over one hour (h), and adjusted for effects of air density variations on CO_2/H_2O fluxes using the standard Webb–Pearman–Leuning correction factor (Webb, Pearman, and Leuning 1980) via the online flux computation. Net radiation and its downwelling and upwelling longwave and shortwave radiation components were measured using net radiometers (CNR-4, Kipp & Zonen USA Inc., Bohemia, NY, USA). Air temperature and relative humidity were measured using an HMP45C probe (Vaisala, Helsinki, Finland). Precipitation, evapotranspiration, soil water content, and mean soil temperature were measured using an ET107 weather-monitoring station (Campbell Scientific, Inc.).

Table 1 List of Hourly Mean Explanatory Variables Measured by the EC System Used in Temporal ANNs

Symbol	Definition	Unit	Comment
F_c	WPL-corrected net CO_2 flux density	mg m^{-2} s^{-1}	Response variable
NEE	Net ecosystem CO_2 exchange	mg m^{-2} s^{-1}	$F_c \cong$ NEE
LE	WPL-corrected latent heat flux density	W m^{-2}	
H_s	Sensible heat flux density	W m^{-2}	Measured by LI-7500
H_c	Sensible heat flux density	W m^{-2}	Calculated from H_s & LE
u^*	Friction velocity	m s^{-1}	
τ	Momentum flux	kg m^{-1} s^{-2}	
T_{air}	Mean air temperature	°C	
CO_2	Mean CO_2 concentration	mg m^{-3}	
H_2O	Mean water vapor density	g m^{-3}	Measured by LI-7500
P_{mean}	Mean air pressure	kPa	
RH	Mean relative humidity	%	
H_2O_{hmp}	Mean water vapor density	g m^{-3}	Measured by HMP45C
ρ	Mean air density	kg m^{-3}	
WD	Wind direction	degrees	
WS	Wind speed	m s^{-1}	
R_n	Mean net radiation	W m^{-2}	
$R_{s\_downwelling}$	Mean downwelling shortwave radiation	W m^{-2}	Not used for nocturnal F_c
$R_{s\_upwelling}$	Mean upwelling shortwave radiation	W m^{-2}	Not used for nocturnal F_c
$R_{l\_downwelling}$	Mean downwelling longwave radiation	W m^{-2}	
$R_{l\_upwelling}$	Mean upwelling longwave radiation	W m^{-2}	
ET	Evapotranspiration	mm	Measured by ET107
T_{soil}	Mean soil temperature	°C	
SWC	Mean soil water content	%	
DOY	Day of year	1 to 365(6)	
h	Hour	1 to 24	

Positive and negative F_c or NEE values denote net CO_2 fluxes to the atmosphere (carbon source) or to the vegetation (carbon sink), respectively.

Data processing

Data processing for hourly mean values of nongap-filled data consisted of two stages: (1) initial removal of missing values or spikes from the data set prior to DWT denoising, a common practice followed by the EC community, and (2) DWT denoising. First, the removal of spikes (erroneous values) was based on the following criteria: (1) the minimum and maximum limits of F_c, H_s (sensible heat), LE (latent heat), u^* (friction velocity), and τ (momentum flux) accepted in the related literature (Olson et al. 2004; Thomas et al. 2011); (2) hours of precipitation events; and (3) a lack of closure of surface energy balance during daytime. The values that fall outside of the minimum and maximum limits were removed when $F_c > 1.76$ or < -1.76 mg m^{-2} s^{-1} (where 1 μmol m^{-2} s^{-1} ≈ 0.044 mg m^{-2} s^{-1}), $H_s > 700$ or < -300 W m^{-2}, LE > 700 or <–100 W m^{-2}, $u^* > 6$ or <0 m s^{-1}, and τ > 100 or <–100 kg m^{-1} s^{-2}

(Olson et al. 2004; Perez-Quezada et al. 2007; Thomas et al. 2011). Likewise, the diurnal values that did not meet the condition of surface energy balance closure were removed using equations (1) and (2):

$$R_n = H_s + \mathrm{LE} + \mathrm{G} + S + \varepsilon \quad (1)$$

$$1.2 * R_n = H_s + \mathrm{LE} \quad (2)$$

where R_n denotes net radiation, G denotes soil surface heat conduction flux, S denotes heat storage in the canopy, ε denotes residual error flux, and 1.2 is a coefficient used to denote the missing sum of G and S not measured in this study. All the negative values ($1.2*R_n - H_s - \mathrm{LE} < 0$) that did not meet the closure quantified using equation (2) were removed (Twine et al. 2000; Xiao et al. 2011).

Second, DWT denoising with the orthogonal wavelet families of Coiflet (coif10), Daubechies (db10), and Symmlet (sym10) with an order of 10 was applied to decompose separately diurnal, nocturnal, and diel F_c data. The best order and decomposition levels for each basis were chosen by trial and error that compared independent validation-derived coefficient of determination (R^2) values obtained when F_c data denoised with different decomposition levels input into the temporal ANNs. The study used the orthogonal wavelet functions because of their unique ability to decompose and reconstruct any signal of time-varying nature without information loss (Mallat 1999; Carl and Kühn 2010).

Unlike the Fourier transform, the DWT converts original time series from the time domain into the concurrent time/frequency domain as father (the low-frequency components) and mother wavelets (the high-frequency components) without loss of temporal information (Kang and Lin 2007; You and Kim 2007; Koirala et al. 2010; Kisi 2011). The three functions of DWT denoising were tested to determine which basis improves temporal ANN models of F_c. KyPlot 2.0 (Kyence Lab. Inc., Tokyo, Japan) was used to estimate these three functions applied to diurnal, nocturnal and diel F_c, and the best polynomial regression models of hourly and monthly mean F_c. The temporal ANN models adopted in this study (TDNN, TLRN, and RNN) are based on the backpropagation through time (BPTT) algorithm, one hidden layer, and the M, which is a locally adaptive approach that uses a memory term to proceed past local minima and to speed up training time (Haykin 1999). Comparisons between B versus O as well as among TDNN, TLRN, and RNN were also carried out to test differences in accuracy performances of the ANNs used.

The temporal ANNs can be divided into two categories: ANNs using memories with (1) a tapped delay line structure such as TDNN and (2) feedback (recurrent) connections such as TLRN and RNN. Temporal ANNs were adopted and compared in this study because (1) they do not require normality assumptions about underlying population distributions; (2) they are especially robust when inputs are highly correlated or have missing values; (3) system behaviors being modeled are nonlinear; and (4) temporal ANNs can capture the temporal context in data by using an adaptive memory of past time periods to predict the future (called BPTT). Accuracy performance metrics used for the temporal ANN models were RMSE, mean absolute error (MAE), and R^2 based on training, cross-validation and independent validation (testing). The fractional random separation of each of the diurnal, nocturnal, and diel data sets used for the ANNs was 60% for training, 15% for cross-validation, and 25% for independent validation (e.g., Hastie, Tibshirani, and Friedman 2009).[1] All the diurnal and diel ANNs of F_c were developed and compared as a function of the same 24 explanatory variables, while all the nocturnal ANNs were built and compared using the same 22 input variables (Table 1). The temporal ANNs were built

using NeuroSolutions 6.01 (NeuroDimension, Inc. Gainesville, FL, USA). Tukey's (1991) multiple comparisons test following one-way analysis of variance was performed to evaluate differences in hourly and monthly mean values of F_c using Minitab 16.1 (Minitab Inc. State College, PA, USA).

Results and discussion

This section presents and discusses the analytical results of (1) processing and temporal partitioning of diel EC data and (2) the effects of DWT denoising, the Coiflet, Symmlet, and Daubechies orthogonal functions, and temporal partitioning on the predictive power and accuracy metrics of the temporal ANNs.

Despiking and temporal partitioning of diel data

The initial removal of missing values reduces the initial raw data set (n = 12,644) by 23.2%, and further despiking the entire data set (n = 9,706) results in a 28.9% decrease.[2] The resultant data set (n = 6,899) is split into diurnal data (daytime, n = 3,577) between 9:00 and 17:00, and nocturnal data (nighttime, n = 3,322) between 18:00 and 8:00. Tukey's multiple comparisons of hourly mean F_c in terms of local hours shows a significant difference among the three distinct periods: (1) the diurnal period of 9:00–17:00; (2) the night period of 20:00–6:00; and (3) the transition periods of early morning (7:00–8:00) and late evening (18:00–19:00) ($P < 0.001$). No significant difference appears between the transition periods (Fig. 1). The monthly mean F_c value in June reaches the highest carbon sink level and differs significantly from the other months, except for July (Fig. 2) ($P < 0.001$).

Diurnal, nocturnal, and diel ANNs with and without DWT denoising

The use of the slowly changing components of nonlinear F_c time series after their separation from the rapidly changing ones through DWT denoising significantly improves the performance metrics of the temporal ANNs that are not obtainable when the temporal ANN models are built using the original data. The best ANNs without denoising are TLRN-1-O-M for the diurnal data and RNN-1-O-M for the nocturnal and diel data, whereas the best ANNs with denoising are TLRN-1-B-M for the diurnal and diel data and TDNN-1-B-M and RNN-1-O-M for the nocturnal data. The best ANNs without denoising have a range of R^2 values indicating an accounting for 15.93% to 29.67% of the nocturnal and diel variations, respectively, whereas R^2 values of the best ANNs with denoising indicate an accounting of 79.5% to 91.45% of the nocturnal and diurnal variations, respectively.

The RMSE values of the best ANNs without denoising vary between 0.1844 and 0.2392 mg CO_2 m^{-2} s^{-1} for the nocturnal and diel time series, respectively, whereas those of the best ANNs with denoising are in the range 0.0193 to 0.0491 mg CO_2 m^{-2} s^{-1} for the diel and diurnal time series, respectively. The best ANNs without denoising have MAE values ranging from 0.0928 to 0.161 mg CO_2 m^{-2} s^{-1} for the nocturnal and diel data, respectively, whereas the MAE values of the best ANNs with denoising varied between 0.0163 and 0.0392 mg CO_2 m^{-2} s^{-1} for the diel and diurnal data, respectively. All of the best ANNs without denoising adopt the online learning algorithm, whereas all those with denoising involve the batch learning algorithm.

Comparing the temporal ANN models built with the DWT-filtered F_c data indicates that the TLRN-1-B-M, TDNN-1-B-M, and RNN-1-O-M models outperform the others in terms of all three performance metrics used. Both conventional statistical techniques (e.g., multiple

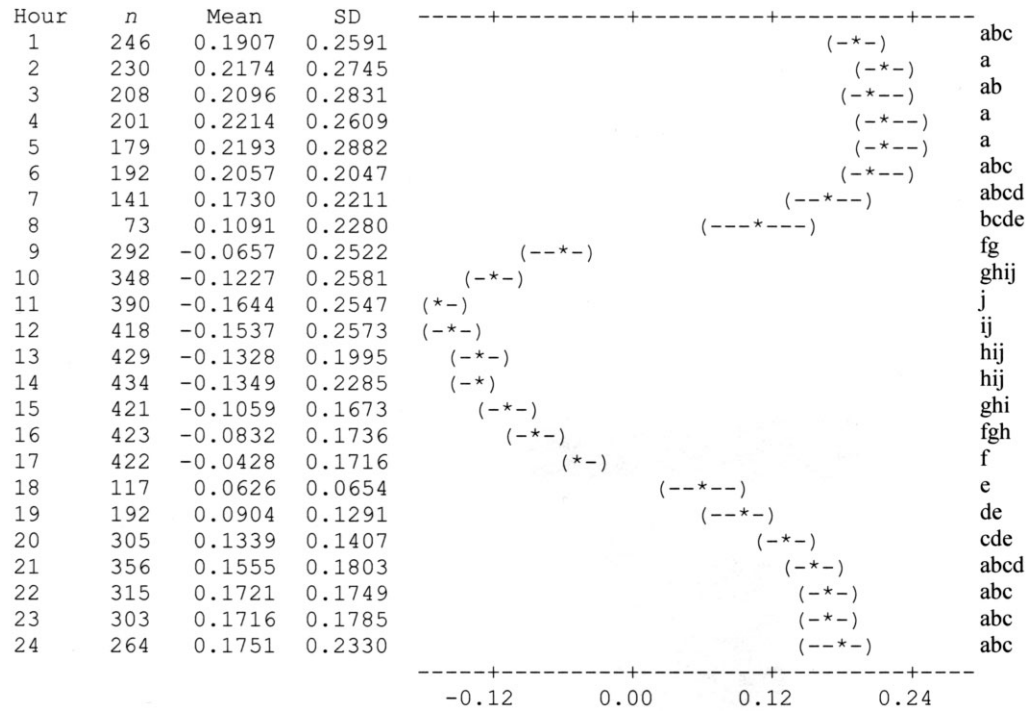

Figure 1. Tukey's multiple comparisons of hourly mean CO_2 fluxes (F_c, mg m^{-2} s^{-1}) ($n = 6899$) between DOY 203 (7 July 2010 11:00) and DOY 719 (12 December 2011 19:00) measured by an EC system for a temperate peatland (Yenicaga, Turkey). Means that do not share the same letter are significantly different at $P < 0.001$. Individual 95% confidence intervals for means display the likely ranges for all the mean differences based on a pooled standard deviation (SD) of 0.2167.

regression) and ANNs suffer from poor generalization capability (performance for unseen data) because of the large number of inputs, their interactions, and (non-) Gaussian data noise. When applying the same methodology used in this study elsewhere (whether it be to rain forest, temperate grasslands, or other peatlands), the generalization ability of ANNs can be developed by (1) training them with different examples, (2) avoiding overfitting in the training, and (3) using cross-validation and independent validation tests (e.g., Gahegan 2000). This study employed these three procedures.

Diurnal, nocturnal, and diel ANNs with different orthogonal wavelet functions

A comparison of the three wavelet functions with the same order of 10 but with different decomposition levels points to the Symmlet with the decomposition levels of eight (sym8) for the diurnal data (Fig. 3), to both the sym7 (Fig. 4) and the Coiflet 7 (coif7) for the nocturnal data, and to the coif9 for the diel data (Fig. 5) as the best performing ones. Unlike the Daubechies (db), the Coiflet and the Symmlet, as nearly symmetrical mother wavelets, maintain the typical shape of the diurnal and seasonal cycles. Temporal partitioning of the diel data set into diurnal versus nocturnal periods improves the testing-derived RMSE and MAE values for the best diurnal and nocturnal ANNs without denoising, and the testing-derived R^2 value for the best diurnal ANN

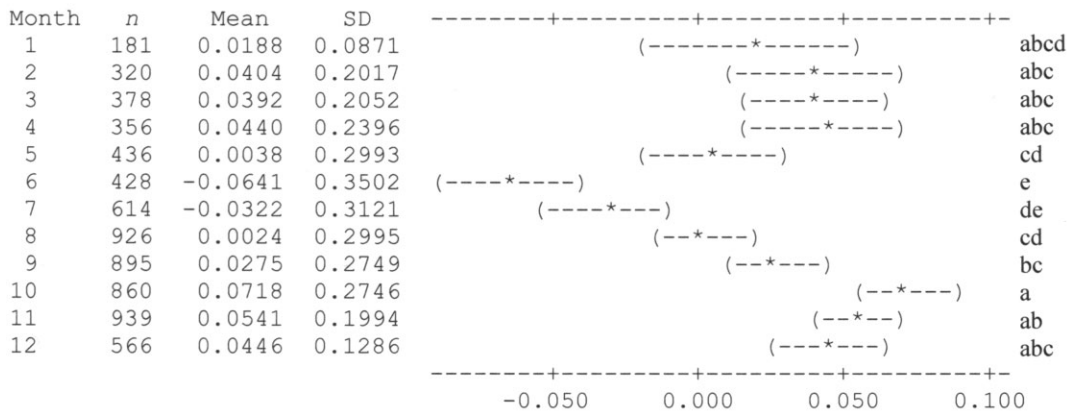

Figure 2. Tukey's multiple comparisons of monthly mean CO_2 fluxes (F_c, mg m^{-2} s^{-1}) ($n = 6899$) between DOY 203 (7 July 2010 11:00) and DOY 719 (12 December 2011 19:00) measured by an EC system for a temperate peatland (Yenicaga, Turkey). Means that do not share the same letter are significantly different at $P < 0.001$. Individual 95% confidence intervals for means display the likely ranges for all the mean differences based on a pooled standard deviation (SD) of 0.2595.

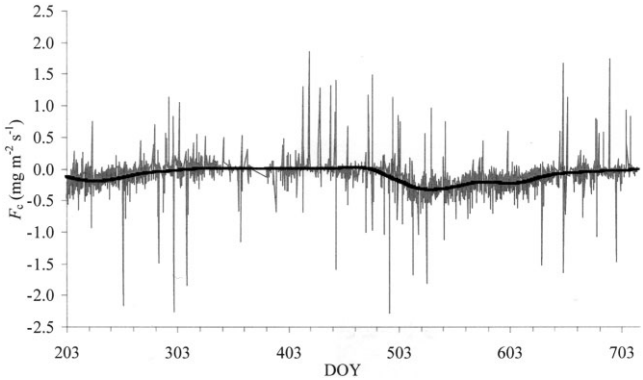

Figure 3. Mean values of diurnal CO_2 flux data (F_c, mg m^{-2} s^{-1}) between DOY 203 (7 July 2010 11:00) and DOY 719 (12 December 2011 19:00) for a temperate peatland (a) without denoising (gray line) and (b) with denoising, based on the best performed orthogonal wavelet basis of Symmlet with an order of 10 and a decomposition level of 8 (black line).

with denoising. The RMSE and MAE values for the best nocturnal ANNs with and without denoising are better than their best diurnal counterparts. The R^2 values for the best diurnal ANNs with and without denoising are better than their best nocturnal counterparts.

ANNs with and without temporal partitioning

Validation of the ANNs without DWT denoising that are subjected to temporal partitioning reveals that the best diel model of RNN-1-O-M has a higher R^2 value by 36–46% than the best diurnal and nocturnal models of TLRN-1-O-M and RNN-1-O-M, respectively. However, temporal partitioning without DWT denoising decreases the RMSE by 33%, and the MAE by 42%,

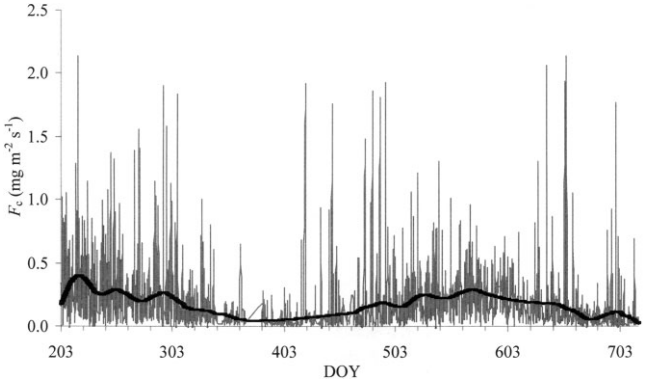

Figure 4. Mean values of nocturnal CO_2 flux data (F_c, mg m^{-2} s^{-1}) between DOY 203 (7 July 2010 11:00) and DOY 719 (12 December 2011 19:00) for a temperate peatland (a) without denoising (gray line) and (b) with denoising, based on the best performed orthogonal wavelet basis of Symmlet with an order of 10 and a decomposition level of 7 (black line).

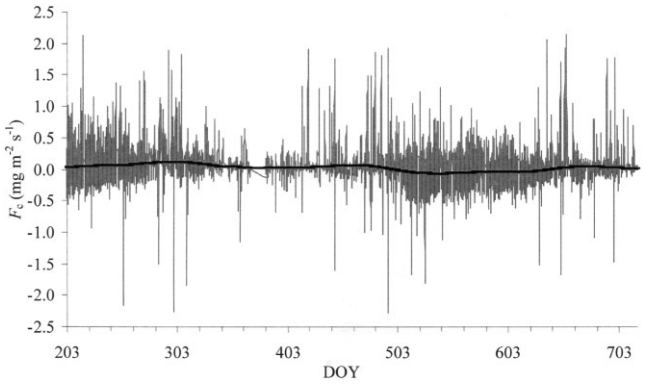

Figure 5. Mean values of diel CO_2 flux data (F_c, mg m^{-2} s^{-1}) between DOY 203 (7 July 2010 11:00) and DOY 719 (12 December 2011 19:00) for a temperate peatland (a) without denoising (gray line) and (b) with denoising, based on the best performed orthogonal wavelet basis of Coiflet with an order of 10 and a decomposition level of 9 (black line).

for the nocturnal ANN relative to the diel ANN (Tables 2–4). The validation-based comparisons of the ANNs with DWT denoising in terms of diurnal or nocturnal versus diel scales point to TLRN-1-B-M as the best diurnal model with the highest R^2 (a 9% increase) and to TLRN-1-B-M as the best diel model with the lowest error measures. The use of both DWT denoising and temporal partitioning markedly improves predictive power and uncertainty measures for the ANNs without DWT denoising, both with and without temporal partitioning. Regardless of the application of DWT denoising, the best diurnal ANNs have higher values of both R^2 and error metrics than the best nocturnal ANNs.

Diurnal ANNs with and without DWT denoising

The diurnal TLRN-1-B-M with sym8-based denoising outperforms the diurnal TLRN-1-O-M without DWT denoising. Relative to the diurnal TLRN-1-O-M without denoising, the diurnal

Table 2 A Comparison of ANN Models of Diurnal CO_2 Fluxes (F_c, mg m^{-2} s^{-1}) between DOY 203 (7 July 2010 11:00) and DOY 719 (12 December 2011 19:00) for a Temperate Peatland with and without DWT Denoising

With and without orthogonal wavelet family-based denoising	Temporal ANNs	Training (n = 2144)			Cross-validation (n = 537)			Independent validation (n = 894)		
		RMSE	R^2 (%)	MAE	RMSE	R^2 (%)	MAE	RMSE	R^2 (%)	MAE
Without denoising	TDNN-1-O-M	0.2663	16.25	0.2062	0.1735	15.06	0.1101	0.2116	6.98	0.1315
	TDNN-1-B-M	0.2159	14.32	0.1215	0.2533	1.06	0.1873	0.2449	10.34	0.1629
	RNN-1-O-M	0.2312	11.99	0.1487	0.2232	7.44	0.1646	0.3237	4.34	0.2555
	RNN-1-B-M	0.2653	7.73	0.1747	0.3753	0.35	0.3007	0.6403	1.63	0.5445
	TLRN-1-O-M	0.2101	21.03	0.1237	0.1750	11.44	0.1096	**0.2056**	**18.90**	**0.1298**
	TLRN-1-B-M	0.2196	15.15	0.1263	0.2578	0.79	0.1887	0.2787	9.22	0.2073
Denoising with sym	TDNN-1-O-M	0.1316	12.91	0.1021	0.0759	7.67	0.0617	0.2187	41.11	0.2058
	TDNN-1-B-M	0.0379	84.85	0.0277	0.0591	2.74	0.0507	0.0580	85.42	0.0464
	RNN-1-O-M	0.0401	84.95	0.0286	0.0716	0.05	0.0577	0.1088	74.13	0.1012
	RNN-1-B-M	0.0373	86.11	0.0292	0.0591	27.76	0.0532	0.0500	90.33	0.0386
	TLRN-1-O-M	0.1780	26.92	0.1584	0.0682	11.69	0.0571	0.2034	47.53	0.1910
	TLRN-1-B-M	0.0272	91.98	0.0197	0.0596	16.28	0.0521	**0.0491**	**91.45**	**0.0392**
Denoising with coif	TDNN-1-O-M	0.1761	25.31	0.1550	0.0758	9.90	0.0622	0.1892	64.13	0.1800
	TDNN-1-B-M	0.0366	85.60	0.0266	0.0769	27.61	0.0674	0.0582	84.98	0.0460
	RNN-1-O-M	0.1385	14.70	0.1119	0.0788	18.46	0.0643	0.2215	20.01	0.2087
	RNN-1-B-M	0.0262	92.68	0.0196	0.0683	28.46	0.0603	0.0712	87.16	0.0594
	TLRN-1-O-M	0.0280	92.66	0.0195	0.0679	4.95	0.0547	**0.0652**	**90.35**	**0.0548**
	TLRN-1-B-M	0.0283	91.37	0.0214	0.0622	15.88	0.0546	0.0722	83.85	0.0573
Denoising with db	TDNN-1-O-M	0.1299	8.02	0.1037	0.0764	11.48	0.0648	0.1912	52.55	0.1812
	TDNN-1-B-M	0.0413	81.16	0.0326	0.0800	29.31	0.0719	0.0519	87.03	0.0394
	RNN-1-O-M	0.0427	83.93	0.0314	0.0729	0.39	0.0607	**0.0708**	**88.36**	**0.0546**
	RNN-1-B-M	0.0305	89.84	0.0230	0.0534	1.74	0.0469	0.0539	85.36	0.0397
	TLRN-1-O-M	0.0970	39.58	0.0726	0.0623	1.04	0.0560	0.1959	48.65	0.1825
	TLRN-1-B-M	0.0248	93.11	0.0198	0.0702	20.27	0.0641	0.0816	88.14	0.0771

Note: based on orthogonal families of Symmlet (sym10-8), Coiflet (coif10-8), and Daubechies (db10-8) with an order of 10 and a decomposition level of 8.

The entries in bold indicate the best performances.

TDNN, time-delay neural network; RNN, recurrent neural network; TLRN, time-lag recurrent network; 1, the number of hidden layers; B, batch learning; O, online learning; and M, Momentum algorithm.

Table 3 A Comparison of ANN Models of Nocturnal CO_2 Fluxes (F_c, mg m^{-2} s^{-1}) between DOY 203 (7 July 2010 11:00) and DOY 719 (12 December 2011 19:00) for a Temperate Peatland with and without DWT Denoising

With and without orthogonal wavelet family-based denoising	Temporal ANNs	Training (n = 1992)			Cross-validation (n = 498)			Independent validation (n = 830)		
		RMSE	R^2 (%)	MAE	RMSE	R^2 (%)	MAE	RMSE	R^2 (%)	MAE
Without denoising	TDNN-1-O-M	0.2141	22.58	0.1096	0.1843	4.08	0.1310	0.2085	8.31	0.1001
	TDNN-1-B-M	0.2096	15.57	0.1200	0.2023	2.32	0.1576	0.2081	3.98	0.1137
	RNN-1-O-M	0.1987	25.68	0.1150	0.1895	9.50	0.1466	**0.1844**	**15.93**	**0.0928**
	RNN-1-B-M	0.3296	1.64	0.2319	0.2767	2.84	0.2155	0.4258	1.09	0.3135
	TLRN-1-O-M	0.2280	13.40	0.1216	0.1939	3.83	0.1380	0.2658	0.77	0.1747
	TLRN-1-B-M	0.2096	15.61	0.1189	0.1900	0.15	0.1413	0.1956	6.21	0.0987
Denoising with sym	TDNN-1-O-M	0.0946	3.57	0.0792	0.0309	53.50	0.0244	0.1323	18.42	0.1049
	TDNN-1-B-M	0.0409	81.15	0.0317	0.0387	21.25	0.0313	0.0430	65.15	0.0356
	RNN-1-O-M	0.0748	60.48	0.0578	0.0280	63.68	0.0232	**0.0341**	**77.23**	**0.0279**
	RNN-1-B-M	0.0384	83.68	0.0290	0.0594	16.02	0.0496	0.0497	75.24	0.0412
	TLRN-1-O-M	0.0839	47.70	0.0673	0.0271	58.80	0.0226	0.0425	63.32	0.0331
	TLRN-1-B-M	0.0545	71.43	0.0417	0.0594	4.39	0.0478	0.0512	57.55	0.0411
Denoising with coif	TDNN-1-O-M	0.0857	29.42	0.0739	0.0376	39.76	0.0286	0.0726	8.14	0.0580
	TDNN-1-B-M	0.0359	85.42	0.0266	0.0445	7.80	0.0346	**0.0389**	**79.50**	**0.0334**
	RNN-1-O-M	0.0433	78.81	0.0333	0.0331	38.63	0.0270	0.0533	48.62	0.0416
	RNN-1-B-M	0.0350	86.44	0.0268	0.0430	4.10	0.0352	0.0627	80.41	0.0562
	TLRN-1-O-M	0.1097	13.56	0.0845	0.0384	24.25	0.0322	0.0703	32.60	0.0568
	TLRN-1-B-M	0.0269	91.80	0.0207	0.0362	29.19	0.0303	0.0498	70.35	0.0406
Denoising with db	TDNN-1-O-M	0.0663	78.82	0.0522	0.0809	19.73	0.0734	0.0545	52.85	0.0448
	TDNN-1-B-M	0.0380	84.51	0.0290	0.0328	30.26	0.0272	0.0454	71.31	0.0371
	RNN-1-O-M	0.1251	15.85	0.0959	0.0448	36.22	0.0383	0.1046	0.44	0.0865
	RNN-1-B-M	0.0451	78.42	0.0352	0.0428	14.58	0.0350	**0.0443**	**76.05**	**0.0372**
	TLRN-1-O-M	0.1020	48.65	0.0790	0.0537	17.90	0.0463	0.1132	31.90	0.0919
	TLRN-1-B-M	0.0386	84.01	0.0297	0.0536	1.42	0.0456	0.0493	65.40	0.0430

Note: based on orthogonal families of Symmlet (sym10-7), Coiflet (coif10-7), and Daubechies (db10-7) with an order of 10 and a decomposition level of 7.

The entries in bold indicate the best performances.

TDNN, time-delay neural network; RNN, recurrent neural network; TLRN, time-lag recurrent network; 1, the number of hidden layers; B, batch learning; O, online learning; and M, Momentum algorithm.

Table 4 A Comparison of ANN Models of Diel CO_2 Fluxes (F_c, mg m^{-2} s^{-1}) between DOY 203 (7 July 2010 11:00) and DOY 719 (12 December 2011 19:00) for a Temperate Peatland with and without DWT Denoising

With and without orthogonal wavelet family-based denoising	Temporal ANNs	Training (4136)			Cross-validation (1035)			Independent validation (1725)		
		RMSE	R^2 (%)	MAE	RMSE	R^2 (%)	MAE	RMSE	R^2 (%)	MAE
Without denoising	TDNN-1-O-M	0.3079	28.21	0.2330	0.2201	56.21	0.1478	0.3498	13.41	0.2543
	TDNN-1-B-M	0.2268	19.57	0.1295	0.2677	30.88	0.1951	0.2174	15.73	0.1259
	RNN-1-O-M	0.2866	20.12	0.2242	0.2648	45.88	0.1884	**0.2392**	**29.67**	**0.1610**
	RNN-1-B-M	0.3071	9.12	0.2143	0.3702	10.39	0.2873	0.2975	12.14	0.2135
	TLRN-1-O-M	0.2851	23.51	0.2193	0.1963	62.67	0.1314	0.2880	22.41	0.2209
	TLRN-1-B-M	0.2274	19.65	0.1364	0.3394	16.75	0.2551	0.2960	1.25	0.1938
Denoising with sym	TDNN-1-O-M	0.0448	37.75	0.0366	0.0169	0.11	0.0145	0.0586	2.89	0.0475
	TDNN-1-B-M	0.0144	80.44	0.0117	0.0295	15.32	0.0256	0.0174	75.51	0.0136
	RNN-1-O-M	0.0311	32.17	0.0257	0.0141	0.79	0.0125	0.0423	30.48	0.0313
	RNN-1-B-M	0.0186	67.00	0.0152	0.0465	8.45	0.0436	0.0329	26.32	0.0302
	TLRN-1-O-M	0.0451	28.39	0.0405	0.0189	6.20	0.0161	0.0542	9.85	0.0437
	TLRN-1-B-M	0.0130	83.94	0.0105	0.0298	5.86	0.0267	**0.0193**	**81.26**	**0.0163**
Denoising with coif	TDNN-1-O-M	0.0579	6.51	0.0492	0.0139	7.28	0.0121	0.0591	35.93	0.0489
	TDNN-1-B-M	0.0151	78.68	0.0125	0.0357	2.02	0.0325	0.0183	83.45	0.0151
	RNN-1-O-M	0.0222	55.89	0.0188	0.0147	6.88	0.0127	0.0635	32.83	0.0527
	RNN-1-B-M	0.0179	69.89	0.0146	0.0272	9.09	0.0235	0.0373	26.96	0.0315
	TLRN-1-O-M	0.0251	66.43	0.0199	0.0186	0.30	0.0165	0.0222	75.12	0.0173
	TLRN-1-B-M	0.0133	83.34	0.0110	0.0283	3.43	0.0254	**0.0190**	**83.75**	**0.0152**
Denoising with db	TDNN-1-O-M	0.0626	24.40	0.0565	0.0330	0.79	0.0275	0.0567	2.09	0.0515
	TDNN-1-B-M	0.0136	79.42	0.0107	0.0497	21.07	0.0395	**0.0289**	**73.39**	**0.0228**
	RNN-1-O-M	0.0505	33.87	0.0437	0.0350	0.02	0.0289	0.0577	12.87	0.0529
	RNN-1-B-M	0.0276	33.48	0.0223	0.0481	4.28	0.0387	0.0445	5.24	0.0411
	TLRN-1-O-M	0.0167	72.80	0.0120	0.0316	3.02	0.0273	0.0496	31.55	0.0451
	TLRN-1-B-M	0.0121	83.57	0.0094	0.0416	13.62	0.0329	0.0336	44.47	0.0276

Note: based on orthogonal families of Symmlet (sym10-9), Coiflet (coif10-9), and Daubechies (db10-9) with an order of 10 and a decomposition level of 9.

The entries in bold indicate the best performances.

TDNN, time-delay neural network; RNN, recurrent neural network; TLRN, time-lag recurrent network; 1, the number of hidden layers; B, batch learning; O, online learning; and M, Momentum algorithm.

TLRN-1-B-M using sym8 results in a 4.8-fold increase in the predictive power (R^2) and 4.2- and 3.3-fold decreases in RMSE and MAE, respectively (Table 2). Among the diurnal ANNs augmented with denoising, TLRN-1-B-M has slightly greater accuracy metrics than TLRN-1-O-M and RNN-1-O-M.

Nocturnal ANNs with and without DWT denoising

The use of wavelet denoising with a decomposition level of seven in the nocturnal ANNs produces better results than the best nocturnal ANN without denoising (RNN-1-O-M). Considering the entire set of accuracy metrics, TDNN-1-B-M with coif7 and RNN-1-O-M with sym7 appear to perform similarly and better than the other ANNs with DWT denoising (Table 3). For example, the R^2, RMSE, and MAE values of the RNN-1-O-M model with sym7 are 4.8-fold greater and 5.4-fold and 3.3-fold less than those of the RNN-1-O-M model without denoising, respectively (Table 3).

Diel ANNs with and without DWT denoising

Among the six diel ANNs without denoising, the RNN-1-O-M model performs best across the set of accuracy metrics. However, the performance of the RNN-1-O-M model is inferior to that of the best diel ANNs with denoising. Among the diel ANNs with denoising, the best performance metrics are obtained using the decomposition level of nine. Among the diel ANNs with denoising based on the decomposition level of nine, the TLRN-1-B-M model with coif9 excels in predicting F_c in terms of R^2, RMSE, and MAE (Table 4). Relative to the RNN-1-O-M model without denoising, the TLRN-1-B-M model with denoising has a 2.8-fold increase in R^2 and 12.5- and 10.6-fold decreases in RMSE and MAE, respectively (Table 4).

As is the case with the other conventional modeling methods, generalizability of findings—that is, the extent to which results from training data can be replicated with data that are new to the model—is an important issue in the ANNs. The trade-off between accuracy and generalizability needs to be balanced when selecting an ANN model in that a model that obtains a perfect emulation of given data is clearly not generalizable, and thus, a generalizable ANN model necessitates controls for overfitting and the use of independent validation data that are separate from the data set from which the ANNs learn and are tested for overfitting (Ripley 1996; Haykin 1999). Overfitting to the training data (overtraining) is most likely to lead to a compromise in the generalizability of the ANN models, which in turn may not render algorithms that eliminate the need for particular sensitivity and specificity for uses in specific biogeoclimate zones. In terms of establishing the generalizability and predictive accuracy of the ANNs, the data in the present study were randomly partitioned into training data from which the ANNs learned, cross-validation data with which the ANNs were tested for overfitting, and independent validation data against which the generalizability of ANN results was gauged. Different noise types and levels with a priori knowledge also may require the selection of different wavelet functions for denoising. However, training ANN models with denoised inputs and outputs without a priori knowledge about types and levels of the noise further enhances the generalizability and predictive performance of the models described in this article.

Conclusions

DWT successively decomposes the original nonstationary F_c signals into the high- (rapidly changing) and low- (slowly changing) frequency components, thus quantifying the signal

features in both time and frequency (scale) domains. Because the selection of a proper denoising function can help to minimize noise associated with the higher frequencies of F_c signals, DWT was executed with the Coiflet, Symmlet, and Daubechies orthogonal functions composed of both mother and father wavelets to characterize the high- and low-frequency components, respectively. The temporal ANNs augmented with DWT using Symmlet and Coiflet excel at predicting nonlinear time series of F_c at diurnal, nocturnal, and diel scales when compared with the temporal ANNs without denoising. When the diurnal and nocturnal F_c time series as complex EC-measured signals are individually or jointly subjected to DWT denoising, the predictive power and errors of the TLRN-1-B-M and TDNN-1-B-M (or RNN-1-O-M) models built using the resultant denoised data are substantially improved, respectively. The science of ecosystem modeling calls for the integrated use of data-driven approaches (e.g., [geo]statistical models and ANNs), process-based models (e.g., biogeochemical and climate models), and denoising algorithms (e.g., wavelets filtering) in order to improve and test our understanding of human-induced disturbances of ecosystem metabolism. This article summarizes the best combination of the six temporal ANNs and three orthogonal wavelet families in quantifying CO_2 metabolism for the peatland ecosystem under anthropogenic influences. The integration of wavelet denoising-based ANNs with spatial analysis and scaling of remotely sensed data remains to be explored in future studies in order to provide new insights into, a better understanding of, and an enhanced comprehension of complex interactions of biogeochemical processes at the ecosystem level.

Acknowledgements

I am grateful to the Scientific and Technological Research Council of Turkey (TUBITAK) for funding this project (CAYDAG-COST 109Y186). I would like to thank Professor Daniel A. Griffith and four anonymous reviewers for their invaluable comments, which significantly improved an earlier version of this article.

Notes

1 A random assignment of data to training, cross-validation, and testing data sets used for the ANNs minimizes bias and ensures generalizability, thus furnishing an appropriate inferential basis.
2 Prior to DWT denoising, the removal of erroneous values (despiking) from nongap-filled EC data as a common preprocessing practice adopted by the EC community involves no spike replacement, as outlined in the "Data processing" section. Therefore, no bias is introduced during despiking.

References

Baldocchi, D. D. (2008). " 'Breathing' of the Terrestrial Biosphere: Lessons Learned from a Global Network of Carbon Dioxide Flux Measurement Systems." *Australian Journal of Botany* 56, 1–26.

Braswell, B. H., W. J. Sacks, E. Linder, and D. S. Schimel. (2005). "Estimating Diurnal to Annual Ecosystem Parameters by Synthesis of a Carbon Flux Model with Eddy Covariance Net Ecosystem Exchange Observations." *Global Change Biology* 11, 335–55.

Carl, G., and I. Kühn. (2010). "A Wavelet-Based Extension of Generalized Linear Models to Remove the Effect of Spatial Autocorrelation." *Geographical Analysis* 42, 323–37.

Daubechies, I. (1988). "Orthogonal Bases of Compactly Supported Wavelets." *Communications on Pure and Applied Mathematics* 41, 909–96.

Daubechies, I. (1992). *Ten Lectures on Wavelets*. Philadelphia: SIAM.

Falge, E., D. Baldocchi, R. J. Olson, P. Anthoni, M. Aubinet, C. Bernhofer, G. Burba, R. Ceulemans, R. Clement, H. Dolman, A. Granier, P. Gross, T. Grunwald, D. Hollinger, N. O. Jensen, G. Katul, P.

Keronen, A. Kowalski, C. Ta Lai, B. E. Law, T. Meyers, J. Moncrieff, E. Moors, J. W. Munger, K. Pilegaard, U. Rannik, C. Rebmann, A. Suyker, J. Tenhunen, K. Tu, S. Verma, T. Vesala, K. Wilson, and S. Wofsy. (2001). "Gap Filling Strategies for Defensible Annual Sums of Net Ecosystem Exchange." *Agricultural and Forest Meteorology* 107, 43–69.

Fuchs, E., C. Gruber, T. Reitmaier, and B. Sick. (2009). "Processing Short-Term and Long-Term Information with a Combination of Polynomial Approximation Techniques and Time-Delay Neural Networks." *IEEE Transactions on Neural Networks* 20, 1450–62.

Gahegan, M. (2000). "On the Application of Inductive Machine Learning Tools to Geographical Analysis." *Geographical Analysis* 32, 113–39.

Gove, J. H., and D. Y. Hollinger. (2006). "Application of a Dual Unscented Kalman Filter for Simultaneous State and Parameter Estimation in Problems of Surface-Atmosphere Exchange." *Journal of Geophysical Research-Atmospheres* 111, D08S07.

Haar, A. (1910). "Zur Theorie der Orthogonalen Funktionensysteme." *Mathematische Annalen* 69, 331–71.

Hastie, T., R. Tibshirani, and J. Friedman. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd ed. New York: Springer.

Haykin, S. (1999). *Neural Network: A Comprehensive Foundation*. Upper Saddle River, NJ: Prentice Hall.

Hollinger, D. Y., J. Aber, B. Dail, E. A. Davidson, S. M. Goltz, H. Hughes, M. Leclerc, J. T. Lee, A. D. Richardson, C. Rodrigues, N. A. Scott, D. Varier, and J. Walsh. (2004). "Spatial and Temporal Variability in Forest-Atmosphere CO_2 Exchange." *Global Change Biology* 10, 1689–706.

Hui, D., S. Wan, B. Su, G. Katul, R. Monson, and Y. Luo. (2004). "Gap-Filling Missing Data in Eddy Covariance Measurements Using Multiple Imputation (MI) for Annual Estimations." *Agricultural and Forest Meteorology* 121, 93–111.

Jung, M., M. Reichstein, H. A. Margolis, A. Cescatti, A. D. Richardson, M. A. Arain, A. Arneth, C. Bernhofer, D. Bonal, J. Chen, D. Gianelle, N. Gobron, G. Kiely, W. Kutsch, G. Lasslop, B. E. Law, A. Lindroth, L. Merbold, L. Montagnani, E. J. Moors, D. Papale, M. Sottocornola, F. Vaccari, and C. Williams. (2011). "Global Patterns of Land-Atmosphere Fluxes of Carbon Dioxide, Latent Heat, and Sensible Heat Derived from Eddy Covariance, Satellite, and Meteorological Observations." *Journal of Geophysical Research-Biogeosciences* 116, G00J07.

Kaheil, Y. H., E. Rosero, M. K. Gill, M. McKee, and L. A. Bastidas. (2008). "Downscaling and Forecasting of Evapotranspiration Using a Synthetic Model of Wavelets and Support Vector Machines." *IEEE on Transactions Geoscience and Remote Sensing* 46, 2692–707.

Kang, S. J., and H. Lin. (2007). "Wavelet Analysis of Hydrological and Water Quality Signals in an Agricultural Watershed." *Journal of Hydrology* 338, 1–14.

Kisi, O. (2011). "Evapotranspiration Modeling Using a Wavelet Regression Model." *Irrigation Science* 29, 241–52.

Knorr, W., and J. Kattge. (2005). "Inversion of Terrestrial Ecosystem Model Parameter Values against Eddy Covariance Measurements by Monte Carlo Sampling." *Global Change Biology* 11, 1333–51.

Koirala, S. R., R. W. Gentry, P. J. Mulholland, E. Perfect, and J. S. Schwartz. (2010). "Time and Frequency Domain Analyses of High-Frequency Hydrologic and Chloride Data in an East Tennessee Watershed." *Journal of Hydrology* 387, 256–64.

Mallat, S. (1999). *A Wavelet Tour of Signal Processing*. New York: Academic.

Melesse, A. M., and R. S. Hanley. (2005). "Artificial Neural Network Application for Multi-Ecosystem Carbon Flux Simulation." *Ecological Modelling* 189, 305–14.

Moffat, A. M., D. Papale, M. Reichstein, D. Y. Hollinger, A. D. Richardson, A. G. Barr, C. Beckstein, B. H. Braswell, G. Churkina, A. R. Desai, E. Falge, J. H. Gove, M. Heimann, D. Hui, A. J. Jarvis, J. Kattge, A. Noormets, and V. J. Stauch. (2007). "Comprehensive Comparison of Gap-Filling Techniques for Eddy Covariance Net Carbon Fluxes." *Agricultural and Forest Meteorology* 147, 209–32.

Nguyen, G. H., A. Bouzerdoum, and S. L. Phung. (2008). *Efficient Supervised Learning with Reduced Training Exemplars*. International Joint Conference on Neural Networks (IJCNN 2008), Hong Kong, 1–6 June 2008, pp. 2981–87.

Olson, R. J., S. K. Hollady, R. B. Cook, E. Falge, D. Baldocchi, and L. Gu. (2004). *Fluxnet: Database of Fluxes, Site Characteristics, and Flux-Community Information*. Oak Ridge: Oak Ridge National Laboratory. ORNL/TM-2003/204.

Ooba, M., T. Hirano, J. I. Mogami, R. Hirata, and Y. Fujinumba. (2006). "Comparisons of Gap-Filling Methods for Carbon Flux Dataset: A Combination of a Genetic Algorithm and an Artificial Neural Network." *Ecological Modelling* 198, 473–86.

Oza, N. C. (2005). "Online Bagging and Boosting." *IEEE International Conference on Systems, Man and Cybernetics* 3, 2340–45.

Papale, D., and R. Valentini. (2003). "A New Assessment of European Forests Carbon Exchanges by Eddy Fluxes and Artificial Neural Network Spatialization." *Global Change Biology* 9, 525–35.

Perez-Quezada, J. F., N. Z. Saliendra, W. E. Emmerich, and E. A. Laca. (2007). "Evaluation of Statistical Protocols for Quality Control of Ecosystem Carbon Dioxide Fluxes." *Journal of the Royal Statistical Society A* 170, 213–30.

Reichstein, M., E. Falge, D. Baldocchi, D. Papale, M. Aubinet, P. Berbigier, C. Bernhofer, N. Buchmann, T. Gilmanov, A. Granier, T. Grunwald, K. Havrankova, H. Ilvesniemi, D. Janous, A. Knohl, T. Laurila, A. Lohila, D. Loustau, G. Matteucci, T. Meyers, F. Miglietta, J. M. Ourcival, J. Pumpanen, S. Rambal, E. Rotenberg, M. Sanz, J. Tenhunen, G. Seufert, F. Vaccari, T. Vesala, D. Yakir, and R. Valentini. (2005). "On the Separation of Net Ecosystem Exchange into Assimilation and Ecosystem Respiration: Review and Improved Algorithm." *Global Change Biology* 11, 1424–39.

Richardson, A. D., M. Mahecha, E. Falge, J. Kattge, A. M. Moffat, D. Papale, M. Reichstein, V. J. Stauch, B. H. Braswell, G. Churkina, B. Kruijt, and D. Y. Hollinger. (2007). "Statistical Properties of Random CO_2 Flux Measurement Uncertainty Inferred from Model Residuals." *Agricultural and Forest Meteorology* 147, 209–32.

Ripley, B. D. (1996). *Pattern Recognition and Neural Networks*. Cambridge: Cambridge University Press.

Saito, M., and J. Asanuma. (2008). "Eddy Covariance Calculation Revisited with Wavelet Cospectra." *SOLA* 4, 49–52.

Sattari, M. T., K. Yurekli, and M. Pal. (2012). "Performance Evaluation of Artificial Neural Network Approaches in Forecasting Reservoir Inflow." *Applied Mathematical Modelling* 36, 2649–57.

Stauch, V. J., A. J. Jarvis, and K. Schulz. (2008). "Estimation of Net Carbon Exchange Using Eddy Covariance CO_2 Flux Observations and a Stochastic Model." *Journal of Geophysical Research-Atmospheres* 113, D03101.

Thomas, M. V., Y. Malhi, K. M. Fenn, J. B. Fisher, M. D. Morecroft, C. R. Lloyd, M. E. Taylor, and D. D. McNeil. (2011). "Carbon Dioxide Fluxes over an Ancient Broadleaved Deciduous Woodland in Southern England." *Biogeosciences* 8, 1595–613.

Tukey, J. W. (1991). "The Philosophy of Multiple Comparisons." *Statistical Science* 6, 100–16.

Twine, T. E., W. P. Kustas, J. M. Norman, D. R. Cook, P. R. Houser, T. P. Meyers, J. H. Prueger, P. J. Starks, and M. L. Wesley. (2000). "Correcting Eddy-Covariance Flux Underestimates over a Grassland." *Agricultural and Forest Meteorology* 103, 279–300.

Webb, E. K., G. I. Pearman, and R. Leuning. (1980). "Correction of Flux Measurements for Density Effects Due to Heat and Water Vapour Transfer." *Quarterly Journal of the Royal Meteorological Society* 106, 85–100.

van Wijk, M. T., and W. Bouten. (1999). "Water and Carbon Fluxes above European Coniferous Forest Modelled with Artificial Neural Networks." *Ecological Modelling* 120, 181–97.

van Wijk, M. T., W. Bouten, and J. M. Verstraten. (2002). "Comparison of Different Modelling Strategies for Simulating Gas Exchange of a Douglas-Fir Forest." *Ecological Modelling* 158, 63–81.

Xiao, X., H. C. Zuo, Q. D. Yang, S. J. Wang, L. J. Wang, J. W. Chen, B. L. Chen, and B. D. Zhang. (2011). "On the Factors Influencing Surface-Layer Energy Balance Closure and Their Seasonal Variability over Semi-Arid Loess Plateau of Northwest China." *Hydrology and Earth System Sciences Discussion* 8, 555–84.

You, J., and T. J. Kim. (2007). "Empirical Analysis of a Travel-Time Forecasting Model." *Geographical Analysis* 39, 397–417.